

Scaling Theory of Entanglement at the Many-Body Localization Transition

Philipp T. Dumitrescu,¹ Romain Vasseur,^{2,3} and Andrew C. Potter¹

¹*Department of Physics, University of Texas at Austin, Austin, TX 78712, USA*

²*Department of Physics, University of California, Berkeley, CA 94720, USA*

³*Materials Science Division, Lawrence Berkeley National Laboratories, Berkeley, CA 94720, USA*

(Dated: January 19, 2017)

We study the universal properties of eigenstate entanglement entropy across the transition between many-body localized (MBL) and thermal phases. We develop an improved real space renormalization group approach that enables numerical simulation of large system sizes and systematic extrapolation to the infinite system size limit. For systems smaller than the correlation length, the average entanglement follows a sub-thermal volume law, whose coefficient is a universal scaling function. Furthermore, the full distribution of entanglement follows a universal scaling form, and exhibits a bimodal structure that produces universal subleading power-law corrections to the leading volume-law. For systems larger than the correlation length, the short interval entanglement exhibits a discontinuous jump across the transition from fully thermal volume-law on the thermal side, to a pure area-law on the MBL side.

Recent experimental advances in synthesizing isolated quantum many-body systems, such as cold-atoms [1–4], trapped ions [5, 6], or impurity spins in solids [7, 8], have raised fundamental questions about the nature of statistical mechanics. Even when decoupled from external sources of dissipation, large interacting quantum systems tend to act as their own heat-baths and reach thermal equilibrium. This behavior is formalized in the Eigenstate Thermalization Hypothesis (ETH) [9, 10]. Generic excited eigenstates of such thermal systems are highly entangled, with the entanglement of a sub-region scaling as the volume of that region (“volume law”). This results in incoherent, classical dynamics at long times. In contrast, strong disorder can dramatically alter this picture by pinning excitations that would otherwise propagate heat and entanglement [11–17]. In such many-body localized (MBL) systems [18–20], generic eigenstates have properties akin to those of ground-states. They exhibit short-range entanglement that scales like the perimeter of the sub-region [17] (“area law”), and have quantum coherent dynamics up to arbitrarily long timescales [21–27], even at high energy densities [17, 26, 28–31].

A transition between MBL and thermal regimes requires a singular rearrangement in the structure of eigenstates from area-law to volume-law entanglement. This many-body (de)localization transition (MBLT) represents an entirely new class of critical phenomena, outside the conventional equilibrium framework of either thermal- or quantum- phase transitions. Developing a systematic theory of this transition promises not only to expand our understanding of possible critical phenomena, but also to yield universal insights into the nature of the proximate MBL and thermal phases.

The eigenstate entanglement entropy can be viewed as a non-equilibrium analog of the thermodynamic free energy for a conventional thermal phase transition, and plays a central role in our conceptual understanding of the MBL and ETH phases. Describing the entangle-

ment across the MBLT requires addressing the challenging combination of disorder, interactions, and dynamics. Consequently, most studies have therefore resorted to fully microscopic simulation methods like exact diagonalization (ED)[15, 16, 32–34]. The exponential complexity of such fully microscopic methods fundamentally limits them to small systems ($\lesssim 30$ sites), preventing them from accurately capturing universal scaling properties. For example, critical exponents computed from ED violate rigorous scaling bounds [35, 36].

A promising alternative is to eschew a microscopic description, which is not required to compute universal scaling properties, and instead develop a coarse grained renormalization group (RG) description of the MBLT. Two related RG approaches [37, 38] have produced a consistent picture of the MBLT (see also [39]). Nonetheless, both approaches rest on ad-hoc albeit plausible heuristics for computing many-body matrix elements. In this paper, we develop an RG scheme building upon the approach of [38], whose steps are rooted in well-established properties of matrix-elements in MBL and thermal systems. Using this RG scheme, we compute the full scaling structure of entanglement across the transition, by simulating large system sizes with a large number ($10^5 - 10^6$) of disorder realizations, that allow systematic extrapolation to the infinite size limit. The resulting scaling properties depart dramatically from those of conventional equilibrium critical points, highlighting the unusual nature of the MBLT.

RG approach – Our RG approach builds a coarse-grained picture of eigenstates by identifying collective many-body resonances that destabilize the MBL phase. Although this approach is not tied to a particular microscopic model, for concreteness, we picture a chain of spinless fermions with Hamiltonian

$$H = \sum_x (-c_x^\dagger c_{x+1} + \text{h.c.} - \mu_x \rho_x + V \rho_x \rho_{x+1}), \quad (1)$$

Here we have set the hopping strength to unity, $\rho_x = c_x^\dagger c_x$ is the fermion density on site x , and μ_x is a random chemical potential drawn from a uniform distribution on $[0, W]$. The non-interacting system ($V = 0$) is Anderson localized with localization length $x_0 \approx 2/\log(1 + W^2)$ [40]. Interactions ($|V| > 0$) can drive multi-particle collective resonances. For weak interactions, $V \ll W$, the system remains MBL and these resonances restructure the local integrals of motion (LIOMs) from weakly dressed single-particle orbitals to local few-body LIOMs [41–44]. For sufficiently strong interactions, MBL breaks down as all degrees of freedom become inter-resonant.

While finding the true resonances is tantamount to solving the many-body Hamiltonian, close to the continuous MBLT, one expects a scale-invariant structure in which resonances are organized hierarchically and can be constructed via an iterative procedure [38]. Since large many-fermion resonances will drive the MBLT, it is natural to consider an effective model in terms of resonant clusters, i.e. groups of inter-resonating single-particle orbitals, characterized only by coarse grained information: the effective bandwidth Λ_i and typical level spacing δ_i .

To characterize cluster interactions, we retain only the typical amplitude, Γ_{ij} , of matrix elements for transitions changing the states of clusters i and j , and compare this to the corresponding typical energy mismatch ΔE_{ij} between those states. For $\Gamma_{ij} \gg \Delta E_{ij}$, the states of i and j will resonantly admix, whereas for $\Gamma_{ij} \ll \Delta E_{ij}$, the clusters will remain decoupled apart from weak virtual dressing. We draw a sharp line between these regimes and define a resonant coupling if $\Gamma_{ij} > \Delta E_{ij}$. The ambiguity of this partition becomes unimportant for the large clusters determining the transition, since both Γ_{ij} and ΔE_{ij} depend exponentially on fluctuating extensive quantities, and are rarely comparable.

The RG procedure for a chain of L sites with periodic boundary conditions proceeds as follows. Initially, each cluster corresponds to a localized single-particle orbital with bandwidth $\Lambda_i = \varepsilon_i \approx \mu_i$ (ε_i the non-interacting single-particle energy), $\Delta E_{ij} = |\mu_i - \mu_j|$, and $\Gamma_{ij} = V(e^{-|i-j|/x_0} + e^{-|i-j-L|/x_0})$. During an RG step, all clusters connected by a path of resonating bonds are merged into a new cluster $\{i\} \rightarrow i'$. The coarse grained parameters of the newly formed cluster are chosen as follows [45]:

$$\begin{aligned} \Lambda_{i'} &= \sqrt{\sum_i \Lambda_i^2 + \sum_{ij} \Gamma_{ij}^2}, \\ \delta_{i'} &= \Lambda_{i'}/(2^{n_{i'}} - 1), \\ \Delta E_{i'j'} &= \delta_{i'}\delta_{j'}/\min(\Lambda_{i'}, \Lambda_{j'}) \end{aligned} \quad (2)$$

where $n_{i'}$ is the number of sites in cluster i' .

The effective inter-cluster couplings are changed according to two distinct rules, locally mirroring MBL or

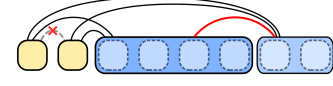


FIG. 1. **Schematic of an RG step.** – Clusters connected by a resonant path ($\Gamma_{ij} > \Delta E_{ij}$) (dashed squares) are merged into bigger clusters. Couplings between clusters are turned off or renormalized (see text).

ETH behavior (Fig. 1). If two clusters are not modified during a RG step, the coupling between them is set to $\Gamma_{i'j'} = 0$ reflecting the emergence of decoupled LIOMs in an MBL system [41, 42]. If at least one of the clusters is modified during the RG step, the new coupling between two clusters is

$$\Gamma_{i'j'} = \left[\max_{i_1 \in \{i\}, i_2 \in \{j\}} \Gamma_{ij} \right] e^{-(n_{i'} + n_{j'} - n_{i_1} - n_{i_2})s_{\text{th}}/2}. \quad (3)$$

This form approximates the resonating clusters as small locally thermal sub-systems with entropy $s_{\text{th}} = \log 2$ per site, as appropriate for matrix elements of local operators in a finite-size, infinite temperature ETH system. We note that a similar form of inter-cluster couplings was recently used to analyze the stability of the MBL phase in various contexts [46].

Intuitively, this procedure describes the state at a given RG step as consisting of a series of fully thermalized or localized clusters. Approximating Γ_{ij} by these limiting forms becomes self-consistently justified since the width of the distribution of resonance parameters $g_{ij} = \Gamma_{ij}/\Delta E_{ij}$ [37, 38, 47] increases with each RG step. In an infinite critical system, the width of the distribution of g increases without bound along the RG flow so that one asymptotically encounters only the cases $g \ll 1$ (MBL) or $g \gg 1$ (ETH) and almost never faces marginal cases where $g \approx 1$. This flow to infinite randomness of g justifies the RG approximations in an analogous fashion to other microscopic RG approaches for quantum phase transitions in disordered spin chains [30, 48–51].

Rooting the Γ_{ij} renormalization rules in well-established asymptotic properties of MBL and ETH systems more accurately captures the competition between the ability for a locally thermal region to act as a bath for its surroundings, versus the tendency for a locally MBL region to isolate neighboring thermal regions. In particular, turning off non-resonant couplings cleanly prevents unphysical “avalanche” instabilities of the MBL phase [46, 52] in which an atypically large resonant cluster becomes increasingly thermal as it grows, enabling it to thermalize an arbitrarily large MBL region [53].

The RG terminates if no resonant bonds remain or the system fully thermalizes. Like [38], our approach allows for a distribution of various cluster sizes in the final configuration. This feature is physically important, as typical configurations at criticality are predominantly

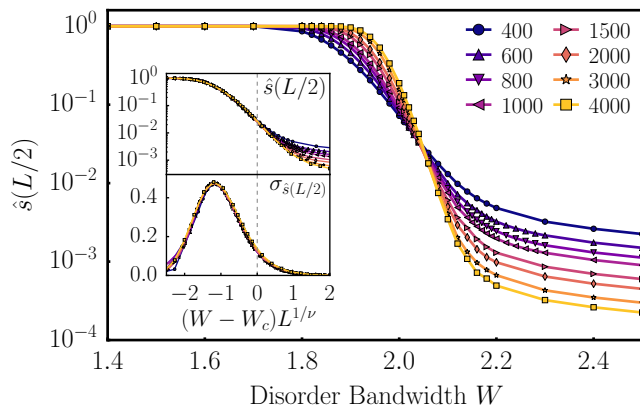


FIG. 2. **Universal scaling of bipartite entanglement** – Normalized bipartite entanglement $\hat{s}(L/2)$ as a function of disorder bandwidth W for different system sizes L . Inset: scaling collapse of $\hat{s}(L/2)$ [upper] and fluctuations $\sigma_{\hat{s}(L/2)}$ [lower], with $W_c = 2.05$ and $\nu = 3.2$. Data with $L \leq 1000$ have at least $2.5 \cdot 10^5$ disorder realizations, and as many as 10^6 near W_c ; those with $L \geq 1500$ at least 10^5 . Error-bars were calculated using Jackknife resampling, but are not shown when smaller than marker sizes.

MBL with few large clusters [38] – a picture supported by recent ED numerics [34]. In contrast, the approach of [37] allowed both MBL and thermal blocks (clusters) to grow until the system is one large block that is either thermal or MBL.

Half-system Entanglement at Criticality – For each disorder realization, the RG produces a configuration of decoupled locally thermal clusters. We calculate the entanglement of a subinterval by summing the thermal volume law contribution from each cluster spanning the interval boundaries. Namely, a cluster partitioned into m and n sites by the entanglement cut contributes $S_{m,n} = \min(m, n)s_{\text{th}}$ [54].

Figure 2 depicts the normalized entanglement entropy

$$\hat{s}(x) = \frac{\overline{S(x, L)}}{xs_{\text{th}}} \quad (4)$$

for $x = L/2$, where $\overline{(\dots)}$ denotes averaging over disorder realizations and interval location. It shows the transition from a fully thermal system consisting of a single large cluster to the localized system made from many small clusters, indicated by a crossing in the curves for different L at a critical disorder $W_c = 2.05 \pm 0.01$. The curves satisfy a scaling form $\hat{s} = f([W - W_c]L^{1/\nu})$, with critical exponent $\nu = 3.2 \pm 0.3$ (Fig. 2 upper inset). This indicates the presence of a single diverging correlation length $\xi \approx |W - W_c|^{-\nu}$. A variety of observables give the same estimates of W_c and ν and our extracted ν lies within error-bars of those obtained in [37, 38]. Notably, we find two distinct values of ν for average and typical correlation length exponents $\nu_{\text{typ}} \approx 2.1 \pm 0.2$ [55], consistent with a flow to infinite randomness. We reiterate that the

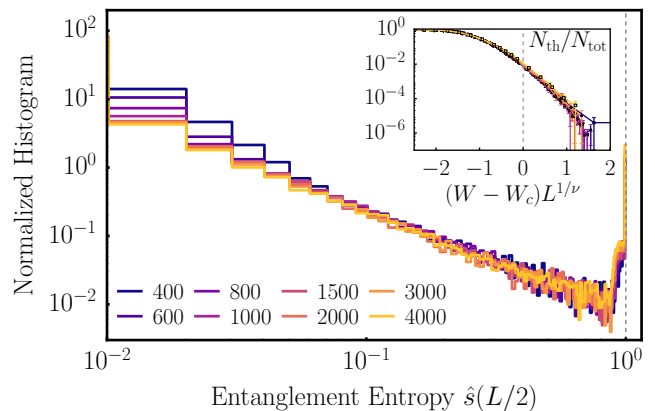


FIG. 3. **Bipartite entanglement at criticality** – Normalized histogram over different disorder realizations of the bipartite entanglement entropy near criticality ($W = 2.04$), using 100 linearly spaced bins. Inset: scaling collapse of fraction of fully thermalized disorder configurations $N_{\text{th}}/N_{\text{tot}}$; error bars are 95% confidence intervals expected for binomial distribution.

small value of the crossing entanglement, and deviation of average and typical exponents demonstrate the transition is driven by rare thermal clusters separated by large MBL regions.

To go beyond the average value, we consider the full distribution of entanglement over disorder realizations. Figure 3 shows the histogram of the normalized entanglement entropy at the MBLT. The distribution has a bimodal structure consisting of a power-law tail, $P(s) \approx s^{-\alpha}$ with $\alpha = 1.4 \pm 0.2$, which we fit over the interval $s \in [0.1, 0.8]$, and a distinct sharp peak near the fully thermal value $s = 1$. Away from criticality, the weight of the thermal peak scales like a universal function of L/ξ , reflected in the fraction of fully thermalizing disorder realizations (Fig. 3 inset). Indications of such a bimodal structure were observed in small-scale ED simulations [56]. Our RG approach allows an extensive exploration of this structure.

At criticality, the thermal peak gives a volume law contribution to the bipartite critical entanglement with a coefficient $a = (0.8 \pm 0.3) \cdot 10^{-2}$ far below the thermal value. The power-law component gives a universal sub-leading contribution to the entanglement intermediate between area- and volume law

$$\overline{S(x = L/2, L, W = W_c)} \approx ax + bx^{1-\alpha} + \dots \quad (5)$$

We note, that these results differ from those of [37], whose proxy for half-system entanglement showed a smaller power-law ($P(s) \approx s^{-0.9}$) and lacked a thermal peak.

Non-local influence of system size – We next investigate the behavior in an infinite system slightly away from the critical point. Near a conventional continuous phase transition, observables (including entanglement) measured over distance x exhibit critical behavior over

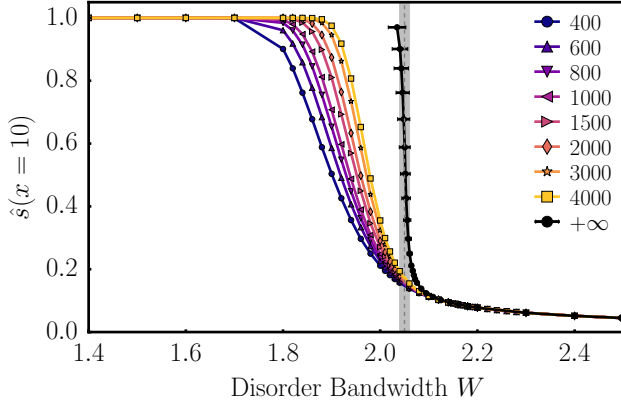


FIG. 4. **Infinite system thermal entanglement** – The normalized entanglement entropy for an interval length $x = 10$ develops a non-analytic step on the thermal side of the MBLT as $L \rightarrow \infty$. Points labeled $+\infty$ are extrapolations in L , assuming the leading scaling form $\propto L^{1/\nu}$ for cuts of fixed \hat{s} , where cubic spline interpolations were used to fill in the curves between data points. The error bars reflect the effect of uncertainty in $\nu = 3.2 \pm 0.3$. The transition $W_c = 2.05 \pm 0.1$ is indicated by the dashed line and gray shaded region.

an extended “critical fan” $x \ll \xi$ extending across both sides of the transition. Moreover, they become independent of system size as $L \rightarrow \infty$, since critical fluctuations are determined by local physics. Entanglement at the MBLT departs dramatically from this conventional behavior, and instead shows a strong non-local dependence on system size, since an infinite thermal system can act as a bath for any finite subsystem, no matter its local properties. Therefore, all subintervals of an infinite system must exhibit fully thermal entanglement $\hat{s}(x, L = \infty) = 1$ for $W < W_c$ and $L \gg \xi$ [57]. The conventional scaling picture would then suggest full thermal entanglement also on the MBL side ($W > W_c$) for $x \ll \xi$ [57]. Instead, ED simulations in [34] give evidence that this region actually has sub-thermal entanglement, consistent with the picture of [38] that the critical regime contains predominately large MBL regions. Together with [57], this implies that the entanglement jumps discontinuously from fully thermal to sub-thermal across the MBLT for $L \gg \xi$ [34].

Our RG approach can directly demonstrate this predicted discontinuity by systematically extrapolating to the limit $L \rightarrow \infty$ with $x \ll \xi \ll L$. Figure 4 shows the normalized entanglement for a fixed interval $x = 10$ and various system sizes L . The black points in the figure represent extrapolations in system size for constant \hat{s} . While one can never observe a true discontinuity in a finite size system, we observe a clear finite size flow towards a non-analytic jump with increasing L . Similar $L \rightarrow \infty$ extrapolations are obtained for all x and the x dependence is negligible on the thermal side of the transition.

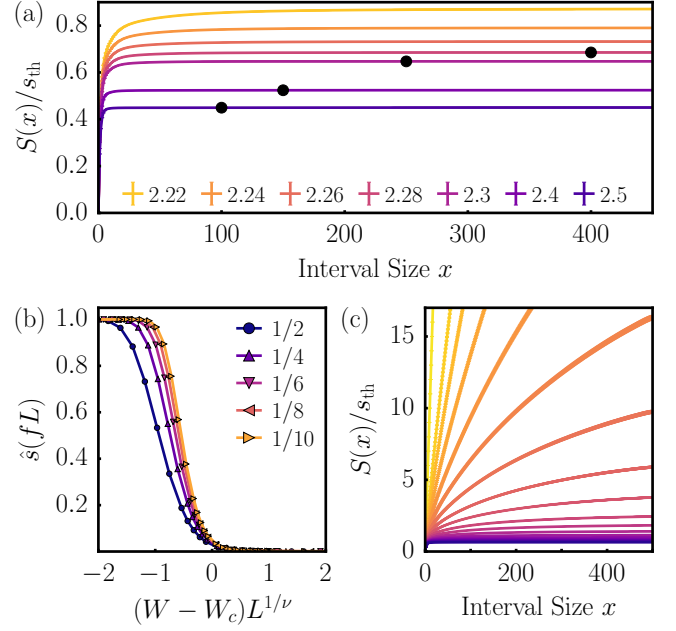


FIG. 5. **Entanglement finite-size crossover** – (a) Entanglement entropy as a function of interval size x for system size $L = 1000$ and various values of $W > W_c$. The error bars correspond to the vertical thickness of the curves. The black points are lower bounds on the estimate of ξ taken from the cluster size histograms. (b) Normalized entanglement entropy for $L = 4000$ for different fractions of system size $f = x/L$. (c) Plot as in (a) zoomed out for disorder values $W = 1.4$ (yellow, linear volume law) and $1.96 \leq W \leq 2.3$ in steps of 0.02.

This discontinuous jump establishes that the entanglement on the MBL side is sub-thermal for all x . However, many functional forms are consistent with this requirement, including sub-thermal volume laws, area-law, or even logarithmic or power-law growth with x . Unlike the thermal behavior for $L = \infty$ and $W < W_c$, which follows from analytic constraints [57], determining the entanglement scaling for $W > W_c$ requires a quantitative exploration of $S(x, L)$ with a three-fold hierarchy of scales $1 \ll x \ll \xi \ll L$ (Fig. 5). This hierarchy requires large systems with at least $\mathcal{O}(10^3)$ sites, making our RG approach uniquely suited to address this question.

Having an objective measure of the correlation length ξ is vital to clearly identify the desired scaling regime and separate it from the distinct crossover behavior when $\xi \approx L$. To this end, we examine the distribution of cluster sizes [55]. Like the bipartite entanglement histograms, they exhibit power-law decay up a scale that we identify as ξ , beyond which they decay exponentially [55]. For $L \gg \xi$, the entanglement curves show a small non-universal rise over $x \lesssim 1 - 10$ and then remain perfectly flat as x crosses through ξ , indicating that the entanglement follows a pure area-law everywhere on the MBL side of the transition, even for $x \ll \xi$.

The results for $L = \infty$ scaling properties of entangle-

ment are succinctly summarized by:

$$\lim_{x \rightarrow \infty} \lim_{L \rightarrow \infty} \hat{s}(x, L) = \begin{cases} 1, & W < W_c \\ 0, & W > W_c. \end{cases} \quad (6)$$

Full scaling form of $S(x, L, W)$ – For infinite systems, we have seen that the entanglement jumps discontinuously at the MBLT. For finite L , this jump becomes a smooth crossover. What universal data can we extract from this crossover? One complication is that entanglement is generically not itself a scaling variable and identifying an appropriate scaling form is not straightforward (e.g. for 1D CFTs one needs to consider the log-derivative $\partial S / \partial \log x$ [58, 59]).

For the MBLT, sub-additivity [57] demands that S increase with both x and L , which is impossible with S being a universal function of x/ξ and L/ξ – physically increasing disorder (decreasing ξ) must decrease entanglement [60]. By performing scaling collapses of $S(x, L)$ for fixed x/L and various W we find evidence that the volume law coefficient is a universal scaling function:

$$\hat{s}(x, L) = \frac{\overline{S(x, L)}}{xs_{\text{th}}} = \mathcal{A}\left(\frac{x}{L}, \frac{L}{\xi} \text{sgn } \delta W\right) + (\dots) \quad (7)$$

where (\dots) indicate subleading corrections in x and L that vanish in the scaling limit $x, L \gg 1$. The explicit dependence on the sign of $\delta W = W - W_c$ is typically absent for ordinary critical points, but allows for the different asymptotic forms $\mathcal{A}(x, y \rightarrow \pm\infty)$ needed to reproduce the discontinuous jump described in the previous section. Figure 5(b) shows cuts of $\mathcal{A}(u, v)$ for various u , illustrating the universal crossover of \hat{s} from its thermal value, 1, in the thermal phase to 0 deep in the MBL phase, which shows a marked dependence on the relative size of the entanglement interval, x , to the system volume, L .

Discussion – The non-local dependence on system size for static eigenstate properties (infinite time-averaged quantities) like entanglement contrasts markedly with the behavior of dynamical measurements, e.g. in quenches from non-stationary states. The discontinuous behavior observed in entanglement for $L \rightarrow \infty$ is special to the former. In contrast, due to the logarithmic causal-cone for dynamics at the MBLT [37, 38], dynamical measurements on timescales $\log t \ll L$ cannot be sensitive to the system size, and will exhibit a more conventional critical scaling fan. This fan is symmetric about W_c and extends up length and time scales $x, \log t \approx \xi$ [37, 38]. The absence of a critical scaling fan in the MBL phase $W > W_c$ is somewhat special to the disorder averaged entanglement itself. The scale invariant power-law tail in the entanglement distribution that develops as $W \rightarrow W_c^+$ decays too rapidly to affect the leading term of average entanglement. However, higher moments of this distribution and other observables, can have universal power

law divergences.

Finally, we remark on an open question of the stability of MBL in higher dimensions. Our RG rules closely mirror the principles used to argue that MBL was unstable to thermal avalanche processes in more than one spatial dimension and for longer than exponential range interactions [46]. Therefore our RG would yield the same conclusions. There are suggestive reasons to suspect the dichotomy between fully MBL and fully thermal behavior may break down [61]. Our current approach does not shed light on this issue, but does suggest that if higher-dimensional MBL does exist, the corresponding MBLT may not have the infinite randomness properties found in 1D.

Acknowledgements – We thank David Huse and Tarun Grover for insightful discussions, and acknowledge the Texas Advanced Computing Center (TACC) at the University of Texas at Austin for providing computational resources. This research was supported by the Quantum Materials program of LBNL (R.V.).

-
- [1] I. Bloch, J. Dalibard, and W. Zwerger, *Rev. Mod. Phys.* **80**, 885 (2008).
 - [2] M. Schreiber, S. S. Hodgman, P. Bordia, H. P. Lüschen, M. H. Fischer, R. Vosk, E. Altman, U. Schneider, and I. Bloch, *Science* **349**, 842 (2015).
 - [3] J.-y. Choi, S. Hild, J. Zeiher, P. Schauß, A. Rubio-Abadal, T. Yefsah, V. Khemani, D. A. Huse, I. Bloch, and C. Gross, *Science* **352**, 1547 (2016).
 - [4] H. P. Lüschen, P. Bordia, S. Scherg, F. Alet, E. Altman, U. Schneider, and I. Bloch, ArXiv e-prints (2016), [arXiv:1612.07173 \[cond-mat.quant-gas\]](#).
 - [5] J. Smith, A. Lee, P. Richerme, B. Neyenhuis, P. W. Hess, P. Hauke, M. Heyl, D. A. Huse, and C. Monroe, *Nat Phys* **12**, 907 (2016).
 - [6] J. Zhang, P. W. Hess, A. Kyprianidis, P. Becker, A. Lee, J. Smith, G. Pagano, I.-D. Potirniche, A. C. Potter, A. Vishwanath, N. Y. Yao, and C. Monroe, ArXiv e-prints (2016), [arXiv:1609.08684 \[quant-ph\]](#).
 - [7] S. Choi, J. Choi, R. Landig, G. Kucsko, H. Zhou, J. Isoya, F. Jelezko, S. Onoda, H. Sumiya, V. Khemani, C. von Keyserlingk, N. Y. Yao, E. Demler, and M. D. Lukin, ArXiv e-prints (2016), [arXiv:1610.08057 \[quant-ph\]](#).
 - [8] K. X. Wei, C. Ramanathan, and P. Cappellaro, ArXiv e-prints (2016), [arXiv:1612.05249 \[cond-mat.dis-nn\]](#).
 - [9] M. Srednicki, *Phys. Rev. E* **50**, 888 (1994).
 - [10] J. M. Deutsch, *Phys. Rev. A* **43**, 2046 (1991).
 - [11] P. W. Anderson, *Phys. Rev.* **109**, 1492 (1958).
 - [12] L. Fleishman and P. Anderson, *Phys. Rev. B* **21**, 2366 (1980).
 - [13] I. Gornyi, A. Mirlin, and D. Polyakov, *Phys. Rev. Lett.* **95**, 206603 (2005).
 - [14] D. Basko, I. Aleiner, and B. Altshuler, *Annals of Physics* **321**, 1126 (2006).
 - [15] V. Oganesyan and D. A. Huse, *Phys. Rev. B* **75**, 155111 (2007).
 - [16] A. Pal and D. A. Huse, *Phys. Rev. B* **82**, 174411 (2010).

- [17] B. Bauer and C. Nayak, *Journal of Statistical Mechanics: Theory and Experiment* **2013**, P09005 (2013).
- [18] R. Nandkishore and D. A. Huse, *Annual Review of Condensed Matter Physics* **6**, 15 (2015).
- [19] E. Altman and R. Vosk, *Annu. Rev. Condens. Matter Phys.* **6**, 383 (2015).
- [20] R. Vasseur and J. E. Moore, *Journal of Statistical Mechanics: Theory and Experiment* **2016**, 064010 (2016).
- [21] M. Znidaric, T. Prosen, and P. Prelovsek, *Phys. Rev. B* **77**, 064426 (2008).
- [22] J. H. Bardarson, F. Pollmann, and J. E. Moore, *Phys. Rev. Lett.* **109**, 017202 (2012).
- [23] M. Serbyn, Z. Papić, and D. A. Abanin, *Phys. Rev. Lett.* **110**, 260601 (2013).
- [24] M. Serbyn, M. Knap, S. Gopalakrishnan, Z. Papić, N. Y. Yao, C. R. Laumann, D. A. Abanin, M. D. Lukin, and E. A. Demler, *Phys. Rev. Lett.* **113**, 147204 (2014).
- [25] R. Vasseur, S. A. Parameswaran, and J. E. Moore, *Phys. Rev. B* **91**, 140202 (2015).
- [26] Y. Bahri, R. Vosk, E. Altman, and A. Vishwanath, *Nat Commun* **6** (2015).
- [27] M. Serbyn, Z. Papić, and D. A. Abanin, *Phys. Rev. B* **90**, 174302 (2014).
- [28] D. A. Huse, R. Nandkishore, V. Oganesyan, A. Pal, and S. L. Sondhi, *Phys. Rev. B* **88**, 014206 (2013).
- [29] A. Chandran, V. Khemani, C. R. Laumann, and S. L. Sondhi, *Phys. Rev. B* **89**, 144201 (2014).
- [30] D. Pekker, G. Refael, E. Altman, E. Demler, and V. Oganesyan, *Phys. Rev. X* **4**, 011052 (2014).
- [31] J. A. Kjäll, J. H. Bardarson, and F. Pollmann, *Phys. Rev. Lett.* **113**, 107204 (2014).
- [32] D. J. Luitz, N. Laflorencie, and F. Alet, *Phys. Rev. B* **91**, 081103 (2015).
- [33] D. J. Luitz, N. Laflorencie, and F. Alet, *Phys. Rev. B* **93**, 060201 (2016).
- [34] V. Khemani, S. Lim, D. Sheng, and D. A. Huse, arXiv preprint arXiv:1607.05756 (2016).
- [35] J. Chayes, L. Chayes, D. Fisher, and T. Spencer, *Phys. Rev. Lett.* **57**, 2999 (1986).
- [36] A. Chandran, C. Laumann, and V. Oganesyan, arXiv preprint arXiv:1509.04285 (2015).
- [37] R. Vosk, D. A. Huse, and E. Altman, *Phys. Rev. X* **5**, 031032 (2015).
- [38] A. C. Potter, R. Vasseur, and S. A. Parameswaran, *Phys. Rev. X* **5**, 031033 (2015).
- [39] L. Zhang, B. Zhao, T. Devakul, and D. A. Huse, *Phys. Rev. B* **93**, 224201 (2016).
- [40] This form smoothly interpolates between the limiting forms obtained by the Born approximation, $x_0 \approx W^{-2}$ ($W \ll 1$) and locator expansion, $x_0 \approx 1/\log W$ ($W \gg 1$).
- [41] M. Serbyn, Z. Papić, and D. A. Abanin, *Phys. Rev. Lett.* **111**, 127201 (2013).
- [42] D. A. Huse, R. Nandkishore, and V. Oganesyan, *Phys. Rev. B* **90**, 174202 (2014).
- [43] V. Ros, M. Müller, and A. Scardicchio, *Nuclear Physics B* **891**, 420 (2015).
- [44] J. Z. Imbrie, *Phys. Rev. Lett.* **117**, 027201 (2016).
- [45] Except when the level spacing of one cluster exceeds the bandwidth of the other, where we instead take: $\Lambda_{i'} \geq \delta_{i'} \geq \Lambda_{j'} \geq \delta_{j'}$, we define $\Delta E_{i'j'} = \max(\delta_{i'} - \Lambda_{j'}, \delta_{j'})$.
- [46] W. De Roeck and F. Huveneers, ArXiv e-prints (2016), arXiv:1608.01815 [cond-mat.dis-nn].
- [47] M. Serbyn, Z. Papić, and D. A. Abanin, *Phys. Rev. X* **5**, 041047 (2015).
- [48] D. S. Fisher, *Phys. Rev. Lett.* **69**, 534 (1992).
- [49] D. S. Fisher, *Phys. Rev. B* **50**, 3799 (1994).
- [50] R. Vosk and E. Altman, *Phys. Rev. Lett.* **112**, 217204 (2014).
- [51] R. Vasseur, A. C. Potter, and S. A. Parameswaran, *Phys. Rev. Lett.* **114**, 217201 (2015).
- [52] K. Agarwal, E. Altman, E. Demler, S. Gopalakrishnan, D. A. Huse, and M. Knap, ArXiv e-prints (2016), arXiv:1611.00770 [cond-mat.dis-nn].
- [53] A more careful re-examination of [38] reveals that such avalanches, while rare, were not completely prevented by the RG rules in that approach, though, this did not quantitatively alter the reported critical properties.
- [54] Here, we omit sub-leading finite-size corrections to ETH [62] and non-universal area law contributions from MBL regions, which only affect the non-universal sub-leading area-law contributions to entanglement.
- [55] See supplemental material.
- [56] X. Yu, D. J. Luitz, and B. K. Clark, *Phys. Rev. B* **94**, 184202 (2016).
- [57] T. Grover, ArXiv e-prints (2014), arXiv:1405.1471 [cond-mat.dis-nn].
- [58] R. C. Myers and A. Singh, *Journal of High Energy Physics* **2012**, 122 (2012).
- [59] H. Liu and M. Mezei, *Journal of High Energy Physics* **2013**, 162 (2013).
- [60] T. Grover, private comm.
- [61] S. Banerjee and E. Altman, ArXiv e-prints (2016), arXiv:1610.04619 [cond-mat.str-el].
- [62] D. N. Page, *Phys. Rev. Lett.* **71**, 1291 (1993).

SUPPLEMENTARY MATERIAL

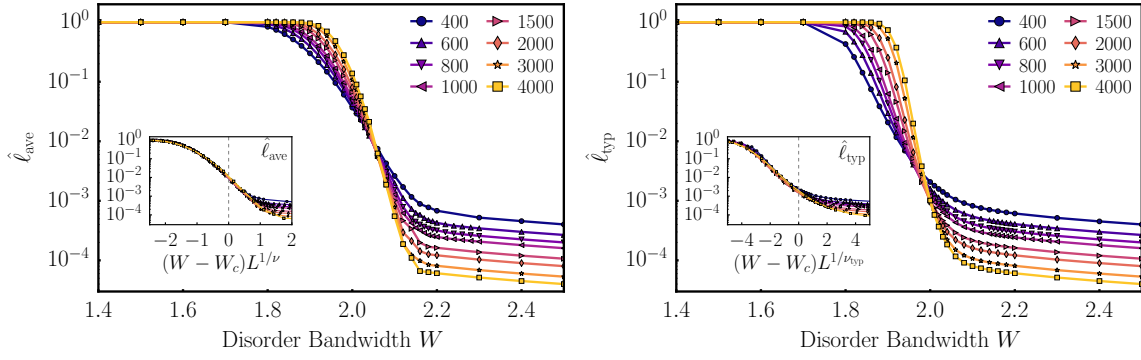


FIG. 6. Normalized average $\hat{\ell}_{ave}$ and typical $\hat{\ell}_{typ}$ cluster length. The typical length $\ell_{typ} = \exp[\overline{\log \ell}]$, where $\overline{[\dots]}$ indicates disorder averaging. The normalization $\hat{\ell}_x = (\ell_x - 1)/(L - 1)$ subtracts the contribution of single site clusters. Insets: scaling collapses with $\nu = 3.2$ and $W_c = 2.05$ [ave] and $\nu_{typ} = 2.1$ and $W_c = 1.99$ [typ]. The typical plot has large finite size corrections and the crossing points drift to higher W ; the typical and average crossings should agree with the true transition as $L \rightarrow \infty$. Nonetheless, the power $\nu_{typ} = 2.1 \pm 0.3$ is consistent with extrapolations along cuts of constant $\hat{\ell}$ away from the crossing point itself. The difference in average and typical scaling is characteristic of a broad distribution and indicates that the transition is rare events driven; this is necessary at an infinite randomness transition.

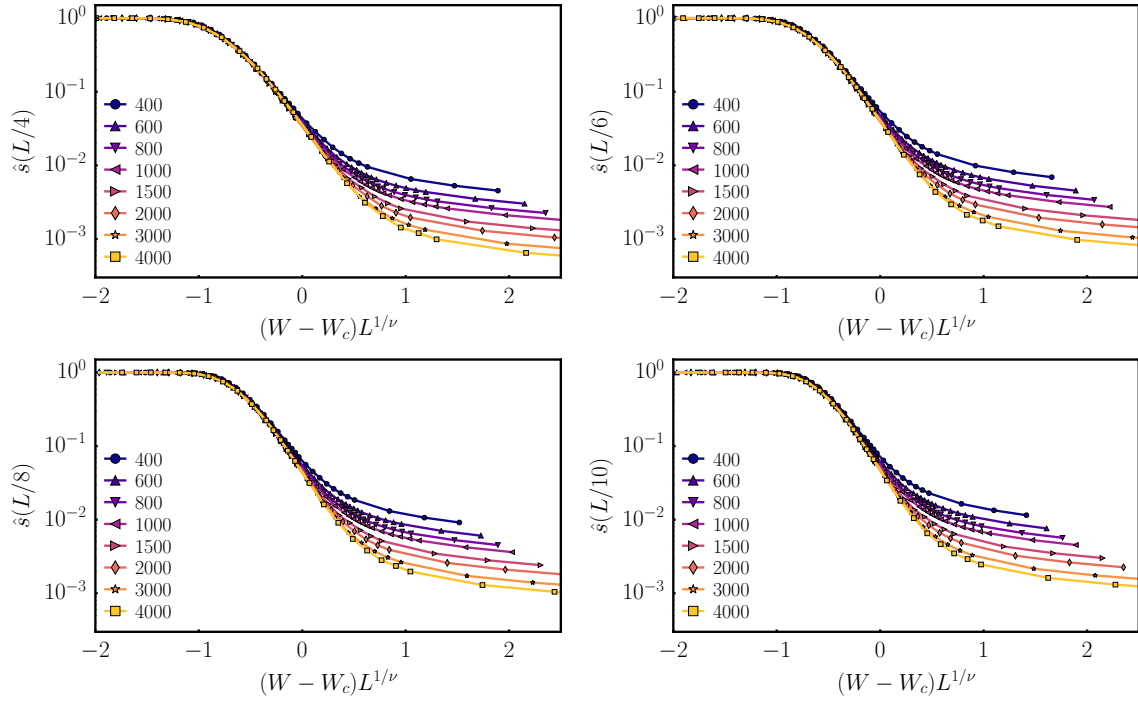


FIG. 7. Scaling collapse of entanglement $\hat{s}(Lf)$ for different system fractions $f = x/L$. Here $W_c = 2.05$ and $\nu = 3.2$, indicating a universal form of the volume law coefficient (7). The $f = 1/2$ scaling collapse is shown in Fig. 2 (inset). As expected, the L curves start separating due to finite size corrections sooner for smaller f . Nonetheless, even for $f = 1/10$, the collapse of $L = 3000, 4000$ persists somewhat for $W > W_c$. The $L = 4000$ data are those shown in Fig. 5(b).

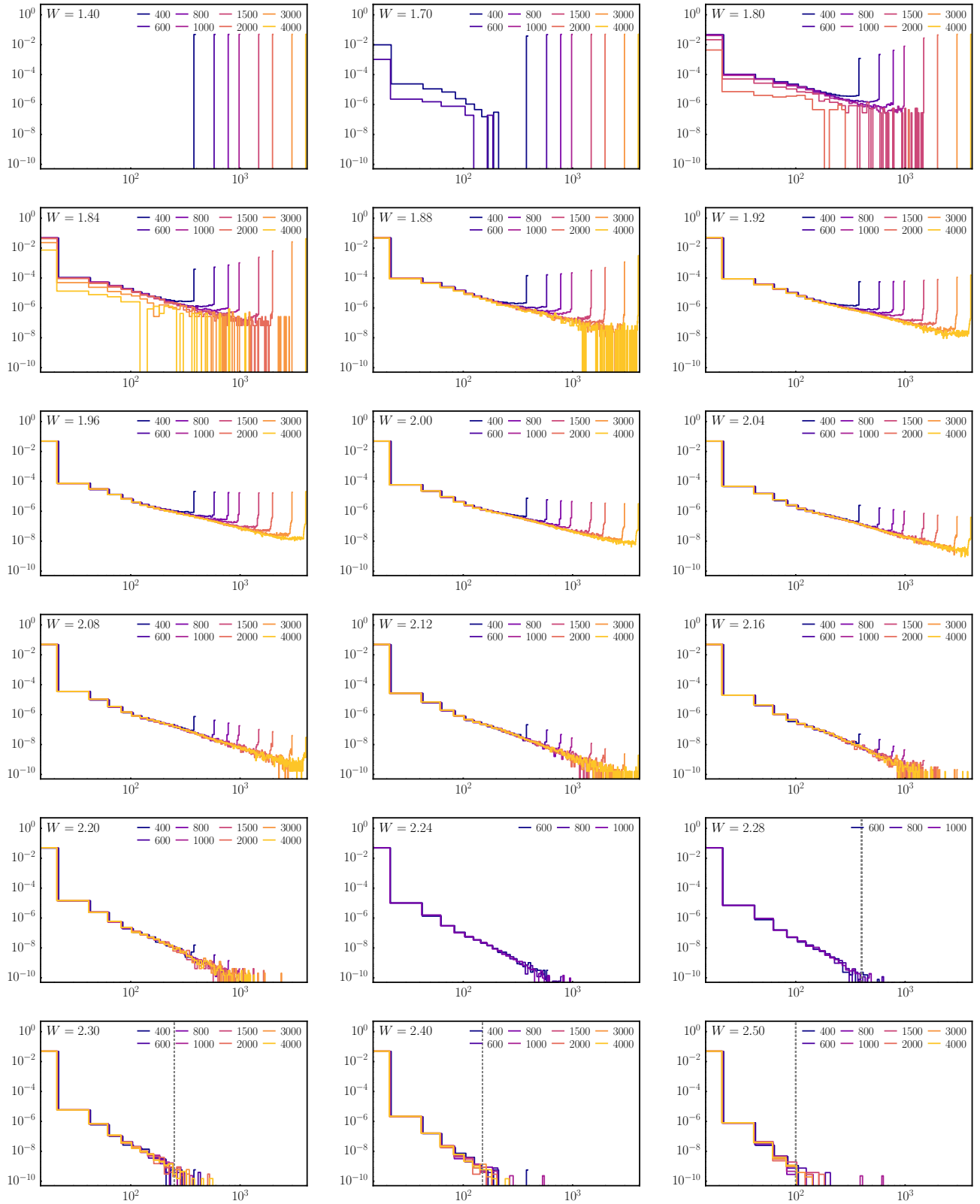


FIG. 8. Normalized histogram over all cluster sizes realized at a given disorder bandwidth W for different system sizes L . Deep in the thermal phase (e.g. $W = 1.4$), each disorder realization consists of one thermal cluster spanning the system. Near the critical point the distribution is bimodal – there is a finite thermal peak and a universal power law. As we move away from the critical point the power law changes to an exponential decay at a length-scale which we can identify as the correlation length ξ . As we care to establish the asymptotic regime $1 \gg x \gg \xi \gg L$, extracting lower bounds on ξ is sufficient. These are given as vertical dashed lines for those W marked in Fig. 5(a). The histograms use $L/20$ bins.